



A new GIS-based data mining technique using an adaptive neuro-fuzzy inference system (ANFIS) and k-fold cross-validation approach for land subsidence susceptibility mapping

Omid Ghorbanzadeh¹ · Hashem Rostamzadeh² · Thomas Blaschke¹ · Khalil Gholaminia³ · Jagannath Aryal^{4,5}

Received: 16 October 2017 / Accepted: 22 May 2018
© The Author(s) 2018

Abstract

In this paper, we evaluate the predictive performance of an adaptive neuro-fuzzy inference system (ANFIS) using six different membership functions (MF). In combination with a geographic information system (GIS), ANFIS was used for land subsidence susceptibility mapping (LSSM) in the Marand plain, northwest Iran. This area is prone to droughts and low groundwater levels and subsequent land subsidence damages. Therefore, a land subsidence inventory database was created from an extensive field survey. Areas of land subsidence or areas showing initial signs of subsidence were used for training, while one-third of inventory database were reserved for testing and validation. The inventory database randomly divided into three different folds of the same size. One of the folds was chosen for testing and validation. Other two folds was used for training. This process repeated for every fold in the inventory dataset. Thereafter, land subsidence related factors, such as hydrological and topographical factors, were prepared as GIS layers. Areas susceptible to land subsidence were then analyzed using the ANFIS approach, and land subsidence susceptibility maps were created, whereby six different MFs were applied. Lastly, the results derived from each MF were validated with those areas of the land subsidence database that were not used for training. Receiver operating characteristics (ROC) curves were drawn for all LSSMs, and the areas under the curves were calculated. The ROC analyses for the six LSSMs yielded very high prediction values for two out of the six methods, namely the difference of DsigMF (0.958) and GaussMF (0.951). The integration of ANFIS and GIS generally led to high LSSM prediction accuracies. This study demonstrated that the choice of training dataset and the MF significantly affects the results.

Keywords Land subsidence susceptibility mapping · Adaptive neuro-fuzzy inference system · k-fold cross-validation · Marand plain · Iran

✉ Omid Ghorbanzadeh
omid.ghorbanzadeh@stud.sbg.ac.at

Extended author information available on the last page of the article

1 Introduction

Land subsidence is the downward motion of a land surface, including rock and soil. This movement happens suddenly or gradually when the underground layers cannot withstand the pressure of the upper layers (Pacheco et al. 2006; Ashraf and Cawood 2015). In most cases, land subsidence is a slow geological process that is hardly noticeable. The final result is a gradual reduction in the total height of the Earth's surface (Ganguli 2011). Factors contributing to this phenomenon include earthquakes and the movement of the earth's crust, melting glaciers, human activities in the fields of drilling and mineral extraction from underground sources, such as oil, and the drainage of water from underground aquifers (Dehghani et al. 2014). More than 80% of land subsidence that occurs in the United States of America has been caused by continuous extraction of groundwater (Galloway et al. 1999). The same situation also applies for Iran's plains in general and the Marand plain in particular. Reductions in the volume of an aquifer through numerous water wells, excavated mainly for agricultural purposes, result in land subsidence. This natural hazard damages natural resources and human properties and constructions. In recent years, several studies have been carried out in the field of land subsidence with different methods and approaches. In southern New Jersey, Sun et al. (1999) used statistical and regression methods to study the relationship between land subsidence and the exploitation of groundwater. In a similar study in West Bengal (India), Ganguli (2011) proposed spreadsheet-based statistical analyses. Lee et al. (2012) assessed ground subsidence susceptibility (GSS) in South Korea using artificial neural networks (ANN). A GIS-MCDA method was also used by Ghorbanzadeh et al. (2017) to identify areas that are prone to land subsidence in the Marand plain, while Dehghan-Soraki et al. (2015) applied ASAR and PALSAR data to analyze and measure the land surface change rates caused by subsidence. Karimzadeh et al. (2013) and Karimzadeh (2015) used the same type of data in order to distinguish the surface deformation in another basin in the Northwest of Iran.

Using modeling to detect areas that are highly susceptible to land subsidence is considered an appropriate way of predicting future possible land subsidence (Lee et al. 2012). Moreover, such predictions are essential for environmental planning managers to control and reduce the adverse impacts of land subsidence (Vaezinejad et al. 2011). Land subsidence susceptibility mapping (LSSM) has been proven as a useful strategy to predict and identify high-risk areas of this phenomenon. Methodologically, there are several approaches for producing the final susceptibility maps. The most common approach is using the spatial models with input data consisting of a variety of related geographic factors and parameters. This is frequently used in environmental modeling (Navas et al. 2012). It has been stated that particularly the availability of final susceptibility maps in combination with a geographic information system (GIS) and efficient database of an area enables the better management of various spatial phenomena (Gaspar et al. 2004). In order to map areas susceptible to natural hazards, some recent attempts have been made to utilize mathematical or statistical methods, such as artificial neural networks (ANN, Lee et al. 2004, 2012; Pradhan and Lee 2010; Chen et al. 2017a, b), fuzzy logic (FL, Pradhan 2011; Feizizadeh et al. 2014a; Shadman Roodposhti et al. 2016), support vector machine (SVM, Feizizadeh et al. 2017), decision tree (DT, Lee and Park 2013), neuro-fuzzy (NF, Vahidnia et al. 2010) and adaptive neuro-fuzzy inference system (ANFIS, Cam and Yildiz 2006; Oh and Pradhan 2011; Bui et al. 2012; Sezer et al. 2011; Basser et al. 2014; Polykretis et al. 2017; Chen et al. 2017a, b). We may conclude from the literature review that a variety of methods has been used to map the susceptibility of the land to environmental hazards. Over the last decade, researchers have

tried to develop more and more accurate methods to map susceptible areas, either through combining different methods such as combining AHP and fuzzy logic (Feizizadeh et al. 2013), using factor score matrix with AHP that called modified AHP (M-AHP) (Nefeslioglu et al. 2013), integrating interval pairwise comparison matrices with the GIS-MCDAs (Feizizadeh and Ghorbanzadeh, 2017) and combining ensemble frequency ratio with logistic regression models (Umar et al. 2014) and analytic network process (ANP) (Pirnazar et al. 2017). There are also several attempts to develop new methods such as the methodology of using the global sensitivity analysis for GIS-MCDAs developed by Ligmann-Zielinska and Jankowski (2012) (Ghorbanzadeh et al. 2017). Our research utilizes the ANFIS by applying the Takagi–Sugeno rule format. This format is a combination of optimized premise MFs with an optimized consequent equation (Bui et al. 2012). ANFIS is an inference system with a high capacity (Sezer et al. 2011). It has also some advantages compared to other susceptibility mapping methods such as expert-knowledge-based GIS-MCDA. GIS-MCDA is sometimes criticized for the expert knowledge to be a major source of uncertainty among the results (Şalap-Ayça and Jankowski 2016; Feizizadeh and Kienberger 2017; Erlacher et al. 2017; Feizizadeh and Ghorbanzadeh 2017; Cabrera-Barona and Ghorbanzadeh 2018) or ordinary neural networks because they use if–then rules (Bardestani et al. 2017). The proposed method does not apply expert opinions at any stage. In addition, it provides the possibility of using a variety of fuzzy MFs. This method is also known for fast convergence times (Pandey and Sinha 2015). By using the input and target data, ANFIS can provide a fuzzy inference structure (Cakit and Karwowski 2017) in which the MF parameters use hybrid learning algorithms to adjust themselves (Bui et al. 2012). The main differences between the literature discussed above and our study are that we used the ANFIS method with different MFs and also k-fold cross-validation (CV) applied to LSSM. We used the k-fold CV approach to deal with the randomness effects on model performance (Gilks and Richardson 1995). With respect to the spatial distribution of land subsidence areas, it is expected that the performance of ANFIS method affect with the random selection of training data. We also used a different approach to bring the layers (maps) related to the ANFIS calculation as input data into the MATLAB software. We used GIS capabilities for the integration of various spatial information with varying spatial resolutions and entities and for applying overlay methods within the ANFIS process. In addition, although the environmental modeling community relies on ANFIS as a powerful and precise method, it has, to our best knowledge, not been used for LSSM. Our study consists of three stages: first, we will introduce the data layers and maps related to land subsidence in the Marand plain. Second, we will produce maps of susceptible areas using these layers combined with hybrid learning and different MFs. Finally, we will validate the maps produced for each of the MFs.

2 Study area

The study area is the Marand plain, which is located in the East Azerbaijan Province in northwestern Iran (Fig. 1). The study area is confined by the mountain ranges of the Mishow Dagħ in the North and the Ghezel-Dagħ and Ghaleh-Dagħ ranges in the South. The northern and southern border areas of the plain to these mountain ranges comprise moderate slopes between 2% and 5%. Most of the precipitation of the study area is snow rather than rain and usually falls during the seasons of autumn, winter and spring with its maximum between December to the end of March. The average annual precipitation is about 242 mm (Fakhri et al. 2015). This study area is located between the Alborz zone to the East and Azerbaijan zone to the North and West, and it belongs to the Central Iran

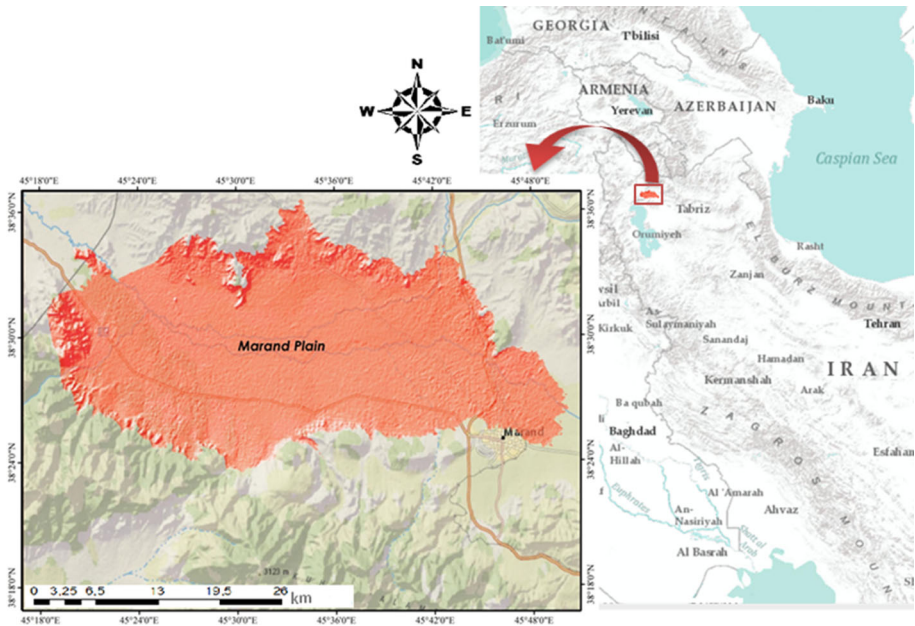


Fig. 1 Study area: Marand plain, northwestern Iran

structural units (Barzegar et al. 2017). The Marand plain is part of the Khazar basin, which is one of the six major hydrological basins in Iran and covers most parts of Northern Iran (Karimzadeh 2015). From a geological point of view, the study area coincides with the same lithological formation, namely a high level piedmont fan and valley terrace deposits Q^{ft1} belonging to the Quaternary age (Khaleghi and Shahverdizadeh 2014).

The primary local activity is agriculture, including irrigated and rain fed agriculture and orchards (Khorrami 2016; Ghorbanzadeh et al. 2017). Agriculture in the region requires a huge amount of surface and underground water. Over the last years, the mean annual precipitation amounts have declined in the study area (Hajalilou and Khaleghi 2009), thus increasing the need to use excessive groundwater extraction (Karimzadeh 2015). This is the main contributing factor leading to the reduction of groundwater levels in this area (Fakhri et al. 2015). Still, the analysis cannot focus on agriculture alone but needs to address complex land use patterns and, in particular, the different water sources such as excavated water wells and dams.

3 Materials and methods

3.1 Data used

In the present LSSM study, seven inter-related factors affect the land subsidence in our study area, including distance to excavated water wells, DEM, slope, land use/cover, depth of groundwater, distance to streams and distance to fault. Ultimately, the lithology layer was not used in the model because almost the entire Marand plain consists of the same lithological formation, namely a high level piedmont fan and valley terrace deposits. All layers were stored in raster format with a pixel size of 100 m (Fig. 2). Although the spatial

resolution of the Landsat satellite images that served as the basis of the land cover map, and the SRTM data on which the topography dataset indicators are based, are 30 m, all pixels were aggregated to 100 m for the modeling for performance reasons and faster computation. As land subsidence phenomena typically cover relatively large areas, this resolution seems appropriate. Moreover, a land subsidence inventory map with a total of 23 land subsidence areas with 522 pixels was prepared to train, test and validate the model.

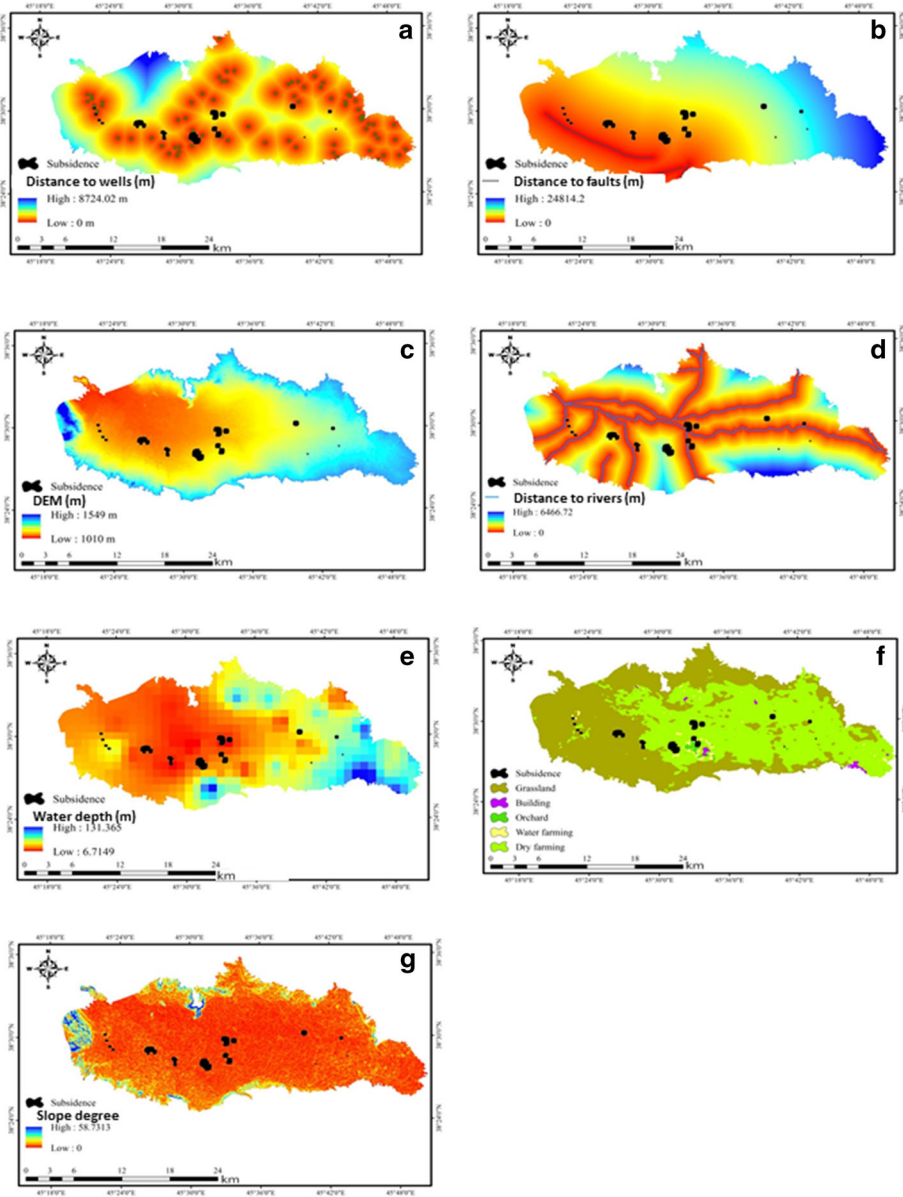


Fig. 2 Seven input land subsidence factors used in the ANFIS model: **a** Distance to wells. **b** Distance to faults. **c** DEM. **d** Distance to rivers. **e** Water depth. **f** Land use/cover. **g** Slope

The land subsidence inventory has been collected by locals based on inquiries from the Department of Natural Resources (East Azerbaijan Province, Iran). 70% (364 pixels) of land subsidence areas within the study area were randomly selected and used as training data. These pixels obtain the information from corresponding pixels of each layer and were used to train the models. The rest of the pixels were used as testing and validation data in two separate groups.

3.2 Land subsidence susceptibility mapping using an adaptive neuro-fuzzy inference system

3.2.1 Fuzzy inference system

A fuzzy inference system (FIS) includes expert knowledge and experience to design a process with input and output fuzzy sets that are controlled by if–then rules (Armaghani et al. 2015a). In simple terms, a FIS is a system which can obtain new knowledge from existing knowledge by using fuzzy logic (Camastra et al. 2015; Cavallaro 2015). A fuzzy inference system is made up of three sections: the first section is the fuzzification process when all crisp values are converted to a linguistic input value using a MF of the system (Tahmasebi and Hezarkhani 2012). The inference engine is the second part and is used to assess the degree of membership of input data based on the output fuzzy sets (Bui et al. 2012). Finally, the fuzzy output values are converted to crisp values in a process called defuzzification (Armaghani et al. 2015a). It can be said that the inference system can produce fuzzy output values based on inference rules as soon as it obtains fuzzy values. The process is presented in Fig. 3.

Generally, three main fuzzy inference systems are used in the literature: The Mamdani model, the Takagi and Sugeno (TKS) model, and the Tsukamoto model. The second model (TKS) is more common (Shabankareh and Hezarkhani 2016) and used in this study. This model renders possible to create fuzzy rules from input data. Moreover, most of the problems do not need rigid conditions in their relative factors which are introduced to the model as input data (Wang et al. 2011). The difference between the Mamdani and TKS models were explained by Cavallaro (2015). The main reason for using the Takagi and Sugeno model in this study is that it is a linear combination of inputs and has fuzzy inputs and crisp outputs (Naderloo et al. 2017). It is also a very efficient computational method for optimization as well as in terms of its implementation (Takagi and Sugeno 1985; Cavallaro 2015).

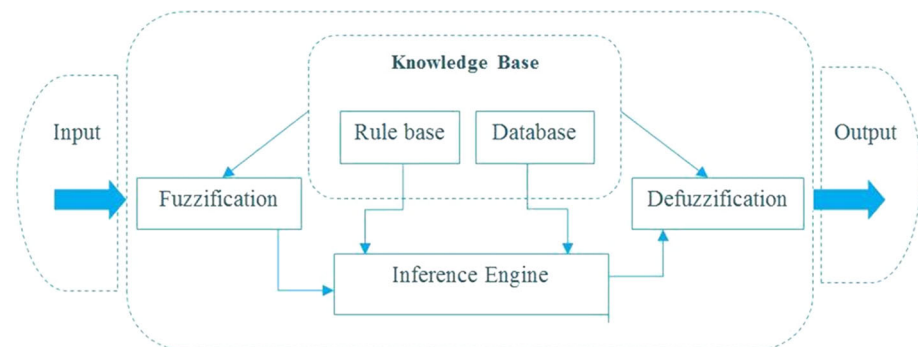


Fig. 3 Structure of a fuzzy inference system

3.2.2 Introducing adaptive neuro-fuzzy inference system structure

ANFIS consists of a hybrid model in which the nodes in different layers of the network provide a neural network for estimating the fuzzy parameters (Polykretis et al. 2017). This model takes advantage of both fuzzy logic and artificial neural networks and combines both approaches making the most of their respective advantages. For further explanation, part *a* of Fig. 4 shows a Sugeno fuzzy model with two rules of fuzzy if-then, with two input values x and y , and f as an output (Jang 1993; Armaghani et al. 2015b).

If x is A_1 and y is B_1 , then $f_1 = p_1x + q_1y + r_1$ (rule 1)

If x is A_2 and y is B_2 , then $f_2 = p_2x + q_2y + r_2$ (rule 2)

The functions of x and y are A_1, A_2, B_1, B_2 and output function parameters include, $p_1, q_1, r_1, p_2, q_2, r_2$. Each fuzzy inference system consists of five different layers and two types of nodes which are adaptive and fixed nodes (see part *b* of Fig. 4).

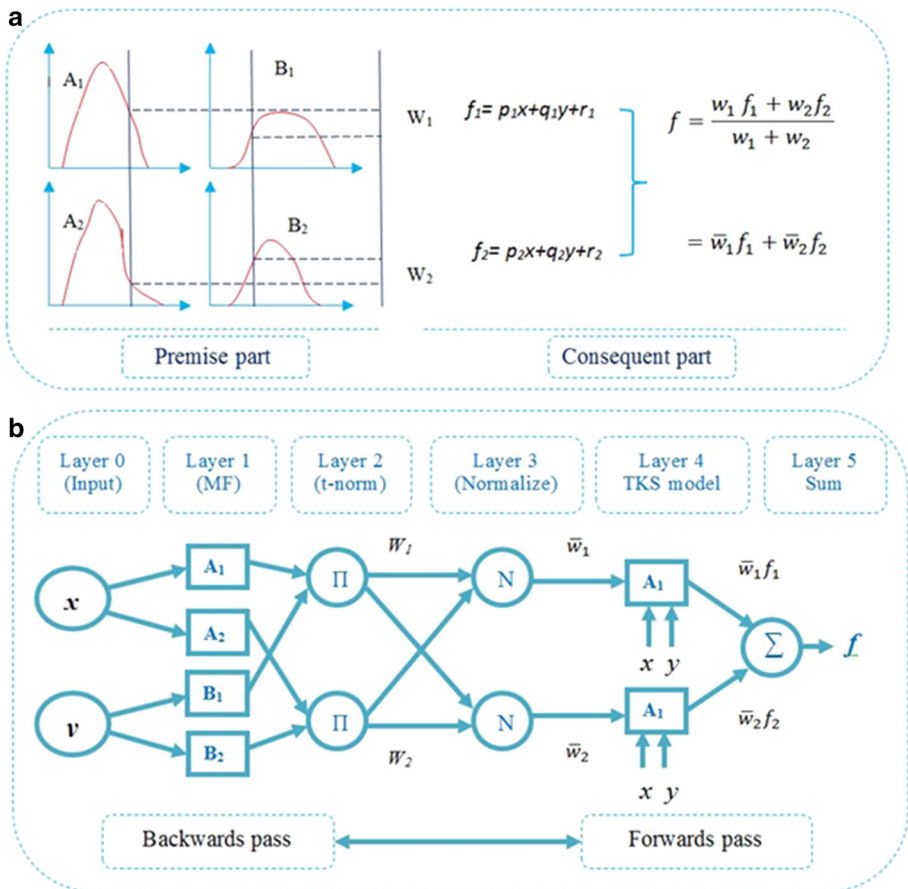


Fig. 4 **a** Sugeno fuzzy model with two rules, **b** typical ANFIS architecture

The difference between these two types of nodes is that adaptive nodes are flexible while fixed nodes are determined through cycles. In the case of differences between layers, it can be noted that the nodes with an index i in layer (1) are adaptive and related to linguistic variables (Lin et al. 2012). Input values are sent directly to the next layer by these nodes.

$$O_{1,i} = \mu A_{i(x)} \quad (1)$$

$$O_{1,i} = \mu A_{i(y)} \quad (2)$$

All nodes in layer (2), which are shown with a Π sign, are fixed nodes, and each node determines the input and output signals (Zhang et al. 2017).

$$O_{2,i} = \omega = \mu A_i(y) B(y) \quad i = 1, 2 \quad (3)$$

In layer (3), all nodes are shown with an N ; they are fixed and are also identified with circles (Bardestani et al. 2017). Another issue in this layer is that outputs are normalized.

$$O_{3,i} = \bar{\omega} = \frac{w_i}{\sum_{j=1}^2 w_j} \quad i = 1, 2 \quad (4)$$

All nodes in layer (4) are adaptive, and their forms are as described by the following function:

$$O_{4,i} = \bar{\omega} f_i = \bar{\omega} (p_i x + q_i y + r_i) \quad (5)$$

Layer (4) is considered to be the final layer. It has only one fixed node which is shown with a \sum representing the sum. The output value of the layer is obtained by the sum of input signals (Basser et al. 2014; Bardestani et al. 2017).

$$o_{5,i} = \sum_i \bar{\omega} f_i = \frac{\sum_i \omega f_i}{\sum_i \omega_i} \quad i = 1, 2 \quad (6)$$

3.2.3 Hybrid learning algorithm

An ANN learning algorithm is used in ANFIS to set up the fuzzy inference system with determined input and output data (Tahmasebi and Hezarkhani 2012). A hybrid learning algorithm was used for training. This algorithm consists of a least-square estimator and gradient descent method (Anwer et al. 2012; Pandey and Sinha 2015). The main objective of the training is to find the optimal parameters for the fuzzy inference system with the minimum value of the error function E , which is the difference between the target amount (t_i) and the output value of the model (f_{out_i}) (Bui et al. 2012). Assume that the training data includes n input data.

$$E = \sum_{i=1}^n (f_{out_i} - t_i)^2 \quad (7)$$

Output that results of layer 5 can be expressed by the equation:

$$f = \bar{\omega}_1 f_1 + \bar{\omega}_2 f_2 = (\omega_1 x) p_1 + (\omega_1 y) q_1 + \bar{\omega}_1 r_1 + (\omega_2 x) p_2 + (\omega_2 y) q_2 + \bar{\omega}_2 r_2 \quad (8)$$

The above equation can be summarized as follows: $f = A.R$ where any of the equation parameters are defined as (9):

$$F = \begin{bmatrix} \text{fout}_1 \\ \text{fout}_2 \\ \dots \\ \text{fout}_n \end{bmatrix}; \quad A = \begin{bmatrix} \bar{\omega}_1 x_1 & \bar{\omega}_1 y_1 & \bar{\omega}_1 & \bar{\omega}_2 x_2 & \bar{\omega}_2 y_2 & \bar{\omega}_2 \\ \bar{\omega}_1 x_2 & \bar{\omega}_1 y_2 & \bar{\omega}_1 & \bar{\omega}_2 x_2 & \bar{\omega}_2 y_2 & \bar{\omega}_2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \bar{\omega}_1 x_n & \bar{\omega}_1 y_n & \bar{\omega}_1 & \bar{\omega}_1 x_n & \bar{\omega}_1 y_n & \bar{\omega}_2 \end{bmatrix}; \quad (9)$$

$$R^T = [p_1 \ q_1 \ r_1 \ p_1 \ q_1 \ r_1]$$

Parameters p_i , q_i , and r_i are unknowns, which can be obtained using the following equation:

$$R = (A^T A)^{-1} A^T f \text{ where } T \text{ is the transpose of the matrix.}$$

The forwards pass of the model moves to the output of the fourth layer of nodes. The backwards pass acts when the next optimal parameters are found (Sengur, 2008; Bui et al., 2012). In this stage, error signals were spreading to the rear side, and premise parameters are restored using gradient descent [see Eq. (10)] (Esen et al. 2017):

$$\Delta \alpha = -\eta \left(\frac{\delta E}{\delta \alpha} \right) \quad (10)$$

where η is a learning rate.

3.2.4 Obtaining land subsidence indexes for mapping

In this study, seven criteria maps were used as input data, and the land subsidence areas layer was defined as the target. The pixels of these areas received all data from the corresponding pixels in the seven criteria maps during the implementation of ANFIS. All these layers were prepared and finally extracted from ArcGIS software (version 10.1) with a pixel size of 509×207 . Each creation layer is composed of 105, 363 pixels in total. The entire ANFIS process, including training and testing, was implemented in MATLAB. Generally, the ANFIS model requires two types of data, namely training data and testing data (Dixon 2005; Bui et al. 2012). There were 23 land subsidence areas in this study, consisting of 522 pixels. k-fold CV approach was used for dividing our inventory dataset into training and testing. In this approach, the dataset D is randomly divide into mutually exclusive k-folds D_1, D_2, \dots, D_k of equal size. Thereafter, the model is run k times and each time $t \in \{1, 2, \dots, k\}$. For the time of t , the model is trained with dataset D without the subset of D_t , and tested with D_t (Kohavi 1995). The selection of the number of folds is depending on several relative factors such as the volume of the inventory dataset, the complexity of the problem and the methodology used. However, in spatial applications, the number of folds is often specified by the researcher without empirical evidence. For instance, Wiens et al. (2008) selected a fivefold cross-validation, Václavík and Meentemeyer (2009) selected a fourfold cross-validation and a threefold cross-validation was chosen by Boria et al. (2014). Considering the size of our database and the amount of computational within six different membership functions, three folds are therefore conducted in this study. As our inventory dataset was divided into three folds, two folds of them served as the training data in each time, while the third fold was used for testing and

validation. In the third fold, pixels were divided randomly into two separate groups, 87 pixels for testing and 87 pixels for data validation. The dispersion of land subsidence inventory in different three folds is shown in Fig. 5. Descriptive information of all layers was transferred to MATLAB as a matrix. The first column of the matrix was the joint field of each pixel for all layers of criteria including the land subsidence areas layer. The joint field is essential to maintain the structure of the maps. It maintains cohesion, arrangements and the position of pixels in all phases of the process. Each of the next seven columns was assigned to descriptive information of each criterion. Finally, the last column was dedicated to land subsidence areas with possible values in this column of zero or one. A value of one indicates that subsidence occurred in that pixel, and a value of zero indicates the absence of the phenomenon. As previously mentioned, the type of fuzzy inference system used is the Sugeno model. It was run with six different types of MFs (see Table 1). Fifty epochs were employed for training the model. The model output for each of the six MFs was a respective matrix with elements between 0 and 1.

Although there are ways of converting the GIS layers into a format that is usable in MATLAB, and vice versa, the fishnet function in ArcGIS was used for converting. In the input data, the value of each pixel was given to a point. Likewise, the output data of the model were transferred into a matrix in the GIS software, while the value of each element was transferred to the corresponding pixels of the fishnet. Pixels that were exactly the same size and had the same number of input layers were placed in the same rows and columns (Parish et al. 2012). As discussed earlier, the LSSM maps were produced using six different MFs. The performance assessments of the trained land subsidence models were examined with two evaluating statistical parameters, namely, the root-mean-square error (RMSE) (11) and the coefficient of determination (R^2) (12), as suggested by various researchers (Bui et al. 2012; Folorunsho et al. 2012; Chai and Draxler 2014; Singh et al. 2017).

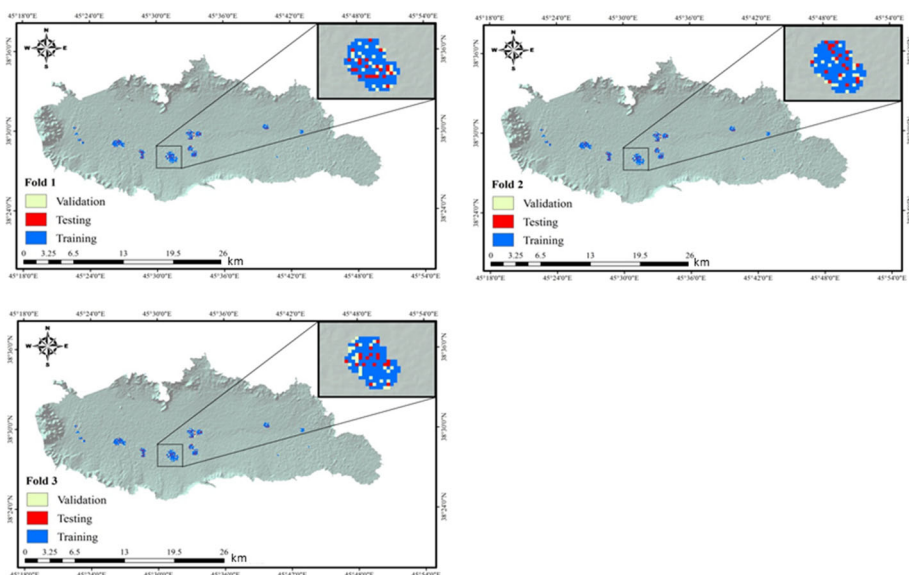


Fig. 5 Varying the dispersion of land subsidence inventory in different three folds

Table 1 Description MFs and statistical criteria of RMSE, R^2 for each fold

No.	Type of MF and descriptions	Fold 1		Fold 2		Fold 3	
		RMSE	R^2	RMSE	R^2	RMSE	R^2
1	Gaussian curve membership function (Gauss)	0.1703	0.912	0.1714	0.963	0.1758	0.950
2	Two-sided Gaussian membership function (Gauss2)	0.1619	0.938	0.1705	0.980	0.1702	0.954
3	Triangular-shaped membership function (Tri)	0.1766	0.917	0.1587	0.967	0.1738	0.933
4	Trapezoidal-shaped membership function (Trap)	0.1825	0.893	0.1773	0.921	0.1827	0.892
5	Difference of two sigmoid membership functions (Dsig)	0.1711	0.952	0.1709	0.951	0.1769	0.947
6	Generalized bell curve membership function (Gbell)	0.1798	0.931	0.1780	0.945	0.1751	0.961

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^n (\text{fout}_i - t_i)^2} \quad (11)$$

$$R^2 = 1 - \left[\frac{\sum_{i=1}^n (\text{fout}_i - t_i)^2}{\sum_{i=1}^n \text{fout}_i^2} \right] \quad (12)$$

where fout_i and t_i are the i th observed and the model predicted outputs (Singh et al. 2012). The observed subsidence polygons were created with GPS from field survey. After that, they mapped in a layer same as other input layers with ArcMap software. Then, all layers transferred to MATLAB as matrices that each element corresponding to a pixel. The pixels within the subsidence polygons were used in developing those matrices and then calculating RMSE and R squared values. In terms of resulted values of R^2 , the best performance was obtained for Gauss2MF in the second fold with the value of 0.980. The best result of RMSE value was 0.158 for TriMF again in the second fold. However, the lowest RMSE and R^2 values were obtained with GbellMF (0.185) and TrapMF (0.892), respectively. Both lowest values were obtained from the third fold of the dataset. All values are shown in Table 1. As we present in Sect. 4, the model got the best result with using the second fold of dataset for testing and rest parts for training. The best results by the second fold are presented in Fig. 6, and other results are not presented keeping in mind the length of the manuscript. Generated maps were classified into five classes from very low to very high susceptibility of land subsidence, and the area of each class is represented in Table 2. They are also compared with each other in Fig. 7.

4 Validation

Validation is considered to be an essential phase in preparing susceptibility maps and in assessing the ability of the model to predict possible future hazards (Pourghasemi et al. 2012; Feizizadeh et al. 2014b). ANFIS method uses training and testing dataset that are randomly selected, the k-fold CV approach was used to mitigate the uncertainties resulted by random selection of dataset (Hahn et al. 2010). The CV approach is often used to select

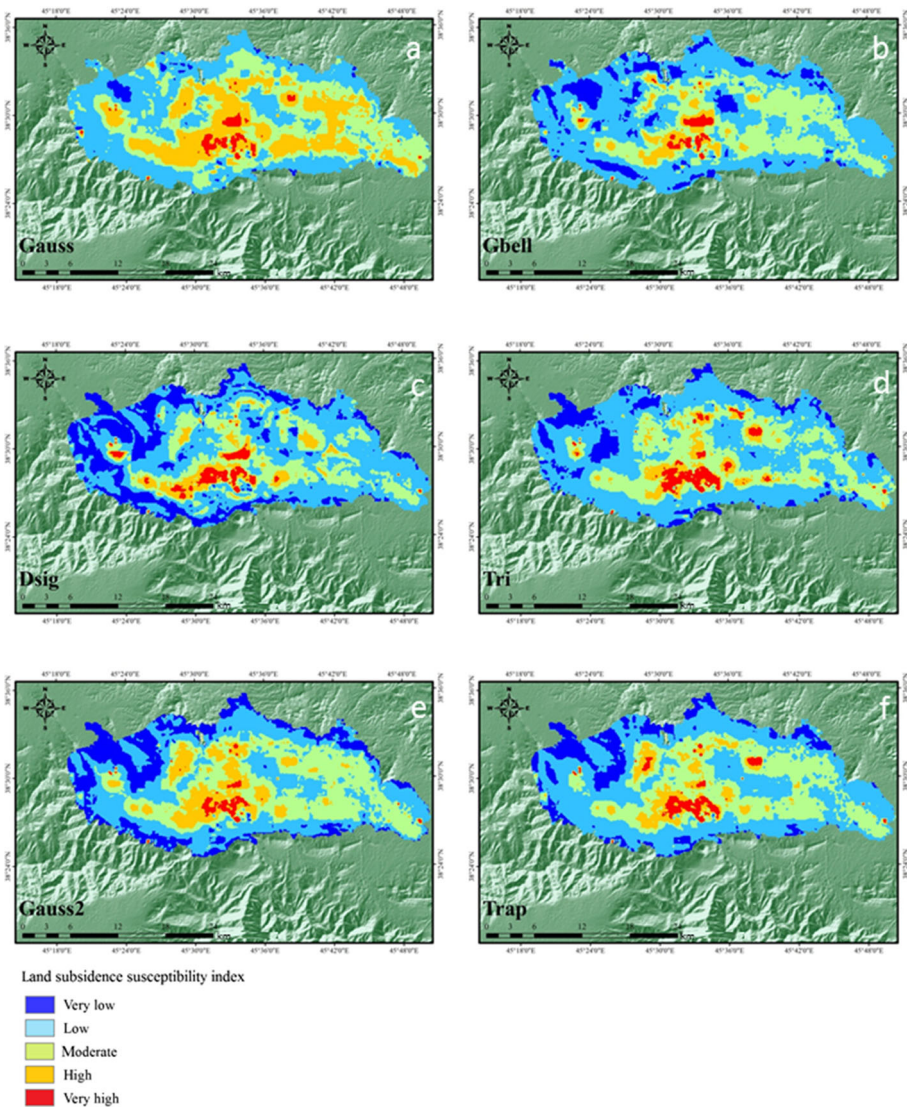


Fig. 6 Land subsidence susceptibility maps using the ANFIS model with the second fold and six different MFs, namely: **a** GaussMF. **b** GbellMF. **c** DsigMF. **d** TriMF. **e** Gauss2MF. **f** TrapMF

the regularization parts in big datasets (Lim and Yu 2016). The simplicity and universality of this approach make it a widespread strategy for the estimate of model performance (Arlot and Celisse 2010). As mentioned in Sect. 3.2.4, we randomly split our dataset D to mutually exclusive 3 folds D_1 , D_2 , D_3 of equal size. Complete CV is calculated by the average of all $\binom{m}{m/k}$ possibilities for any selection of m/k instances out of m (Kohavi 1995). The resulted CV values for each MF are presented in Table 3. In our case, as we have three folds, maximum possibilities are only three. Therefore, the whole process (see Fig. 8) was implemented three times with completely different folds of training and testing

Table 2 The class areas for each of the MFs

Type of MF	Class	Area (ha)	Area (%)	Type of MF	Class	Area (ha)	Area (%)
Gaussian curve				Difference of two sigmoid			
	Very low	1019	1.76		Very low	11870	20.59
	Low	22985	39.88		Low	27628	47.94
	Average	18633	32.33		Average	12610	21.88
	High	13781	23.91		High	3873	6.72
	Very high	1211	2.10		Very high	1648	2.85
Two-sided Gaussian				Trapezoidal-shaped			
	Very low	10739	18.63		Very low	7258	12.59
	Low	21642	37.55		Low	27335	47.43
	Average	17857	30.98		Average	17157	29.77
	High	6462	11.21		High	4347	7.54
	Very high	929	1.61		Very high	1532	2.65
Generalized bell curve				Triangular-shaped			
	Very low	6639	11.52		Very low	6246	10.83
	Low	30536	52.98		Low	29771	51.65
	Average	16178	28.07		Average	15769	27.36
	High	3322	5.76		High	3824	6.63
	Very high	954	1.65		Very high	2019	3.50

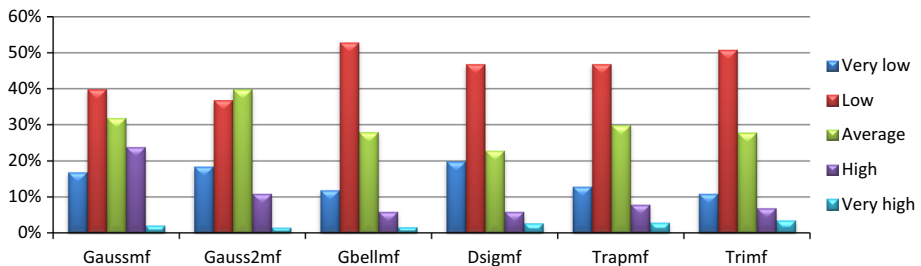

Fig. 7 Comparing the percentage of areas for each of the MFs

Table 3 The percentage of accuracy results of each fold

MF	AUC _{Fold 1}	AVE _{Fold 1}	AUC _{Fold 2}	AVE _{Fold 2*}	AUC _{Fold 3}	AVE _{Fold 3}	CV
Gauss	94.19	92.98	95.12	94.11	95.00	93.79	94.77
Gauss2	91.62		92.79		93.33		92.58
Tri	91.53		92.83		92.22		92.19
Trap	91.46		93.04		92.59		92.36
Dsig	95.08		95.88		95.13		95.36
Gbell	94.01		95.03		94.49		94.51

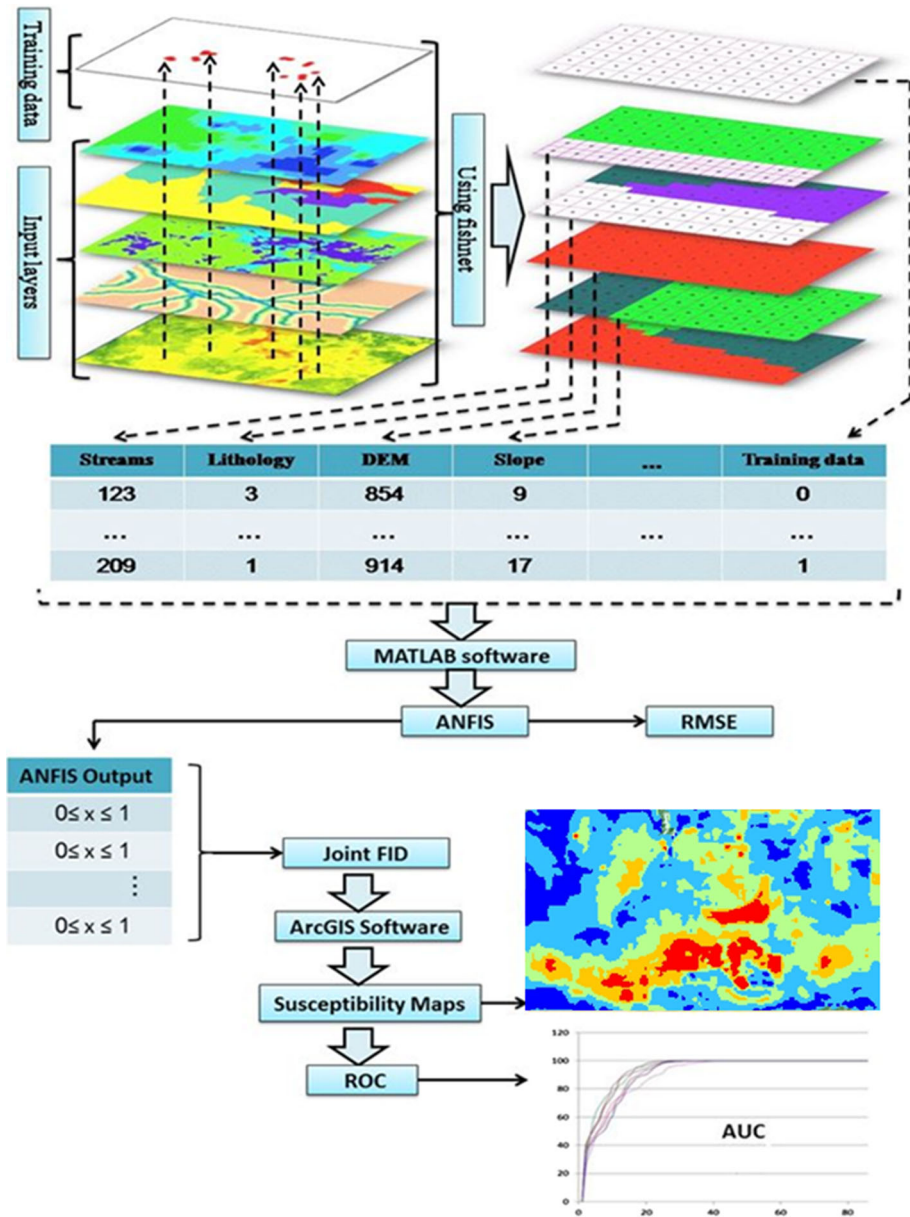


Fig. 8 Flow chart of the production of LSSMs and accuracy assessments for each fold of the inventory database

data. The accuracy of each land subsidence susceptibility map was calculated by applying a common validation procedure based on the known land subsidence areas in the Marand plain. 87 pixels of land subsided areas were used to evaluate the results. Each map was compared to each fold of validation data using the receiver operating characteristics (ROC) curve method. The ROC curve is an efficient method for determining the authenticity of

prediction systems (Swets 1988; Oh and Pradhan 2011) and based on the theory behind the ROC, the area under the curve (AUC) indicates the quality of a prediction system, whereby values close to 1.0 are considered to indicate the best results of a model (Lombardo et al. 2015). Based on resulted LSSMs, ROC curves were calculated separately for each fold (see Fig. 9) and each MF (see Fig. 10). The rate of true-positives was plotted on the vertical axis. The horizontal axis shows the rate of false-positives for all resulting prediction maps. Each scenario corresponds to a point in the plotted space. Finally, the accuracy results of each MF within each fold were calculated and are presented in Table 3. The average values of MF accuracies for each fold are also presented for better comparison. Based on the average values, $AVE_{Fold\ 2}^*$ indicates the most accurate results derived from the ANFIS method using second fold data sets among the three folds. Figure 6 presents the LSSMs that are generated and tested with the second fold dataset. In the case of using this fold model, DsigMF yielded the best result in ROC value (0.958), followed by GaussMF (0.951), GbellMF (0.950), TrapMF (0.930), TriMF (0.928), and Gauss2mf (0.927). DsigMF also got the best result in CV with the ROC value of more than 0.953.

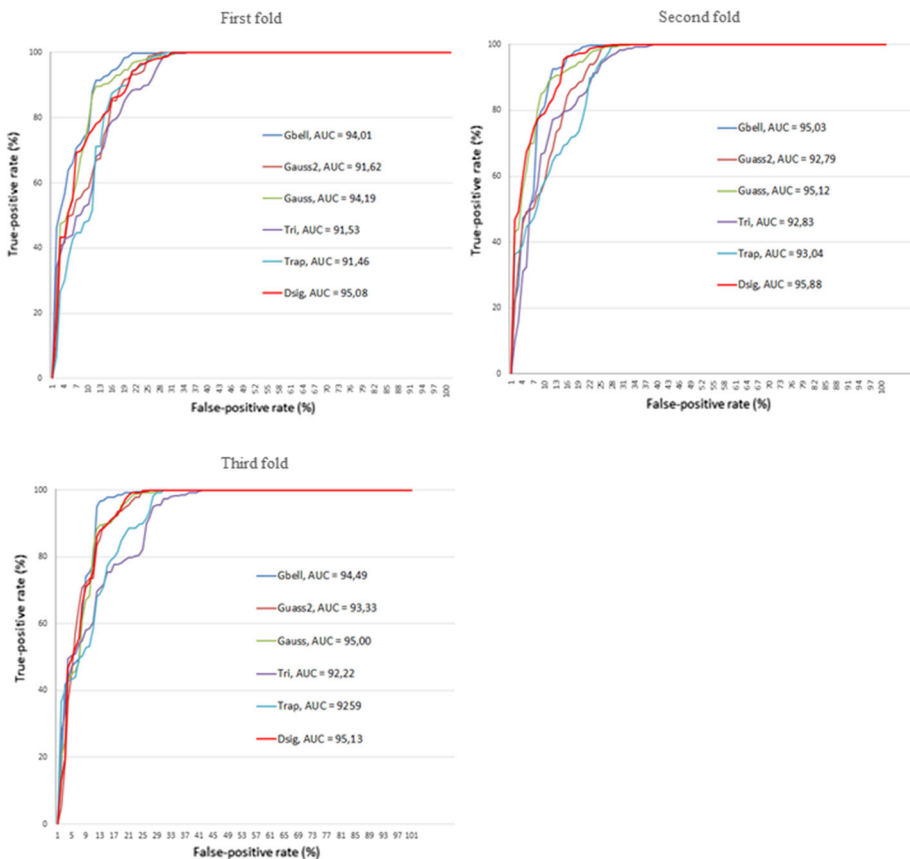


Fig. 9 ROC curves for the resulted maps of each fold

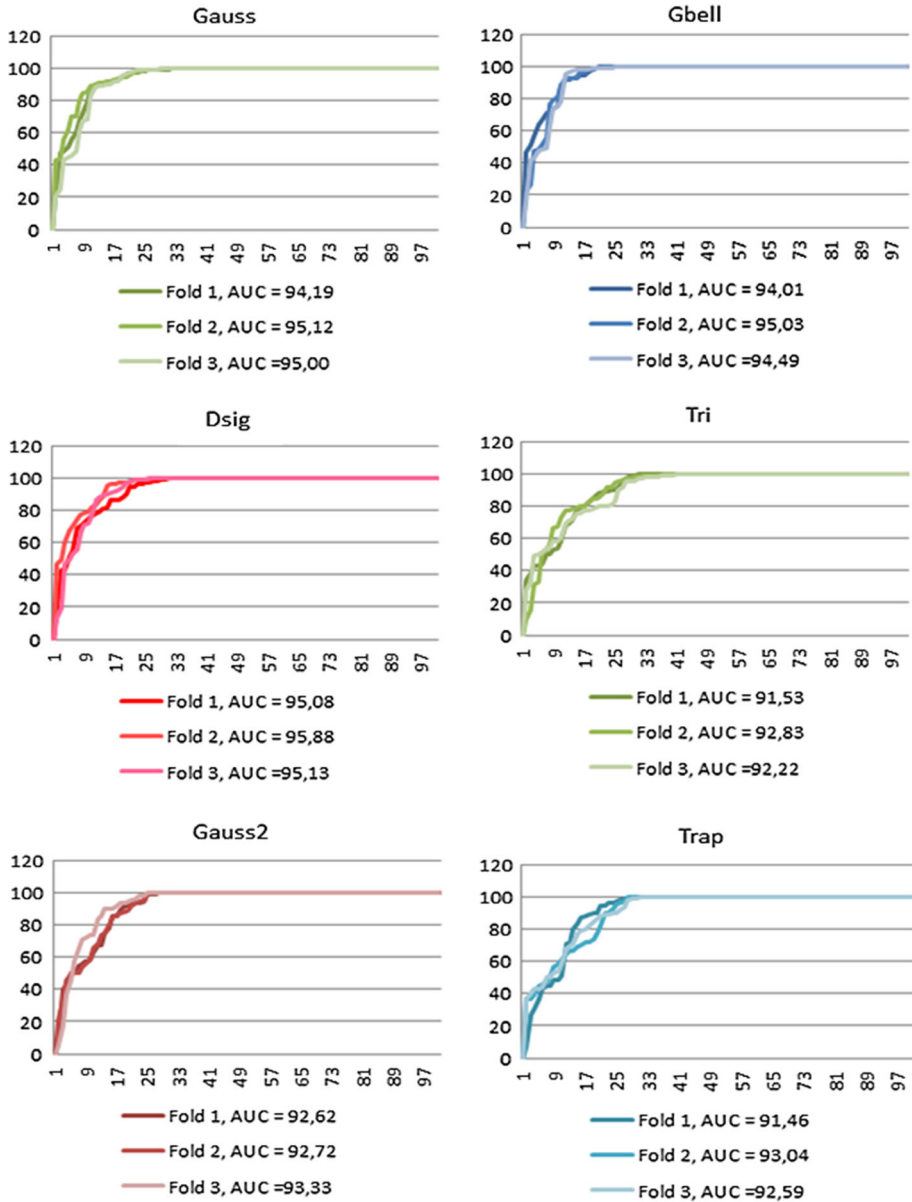


Fig. 10 ROC curves for the resulted maps of each MF

5 Discussion and conclusion

The semi-arid and arid areas of Iran have experienced critical water challenges in the past years. At the same time, the demand for water is increasing due to population growth and more intensive agriculture activities. Therefore, land subsidence occurs in most of the basins of the country because of high ground water extraction (Karimzadeh 2015). Land

subsidence is a severe problem in the plains of northwestern Iran. The study area is particularly susceptible to land subsidence with potentially severe consequences, and the Marand plain is one of six critical plains in the East Azerbaijan Province (MWREP 2014). It may take several years from the onset of land subsidence until visible damages occur at the surface of a plain (Lee et al. 2012). Land subsidence can have a variety of reasons and sometimes a combination of several triggering factors. However, in this study area, massive exploitation of groundwater resources is regarded as the main reason for a severe decline in groundwater levels in the last three decades (Fakhri et al. 2015). In the present study, GIS techniques along with an adaptive neuro-fuzzy inference system (ANFIS) were used for a land subsidence susceptibility assessment in the Marand plain using six different MFs and three folds of dataset. The study consisted of three main steps: the design and structuring of the ANFIS model, the susceptibility mapping, and a validation step for different MFs and folds used. In the first step, an ANFIS model was constructed in MATLAB software environment, and six different MFs were selected to be used in the model. In the second step, seven environmental criteria were identified, including distance to excavated water wells, DEM, land use/cover, distance to fault, depth of groundwater, distance to stream, and slope. These data layers were used as the model inputs. These relative causal criteria played an important role in generating the final land subsidence susceptibility map products. All of the resulting land subsidence susceptibility maps show that the central and western parts of the plain are more susceptible to ground subsidence. Still, as presented in Table 2, there are some differences. The TriMF map revealed the largest area of land highly susceptible to subsidence (3.5%). Gauss2MF yielded the smallest areas with 1.6%. The performance assessment of the ANFIS model revealed Gauss2MF and GbellMF to have the highest performances with 0.98 and 0.96 for R^2 values, and 0.170 and 0.175 for RMSE values in different folds. But this is not to say that Gauss2MF and GbellMF always achieve the best model results. The performance of MFs depends on several parameters, such as the distribution of training and test data, and their performance may differ in other studies with different data. The ROC method was used to validate the six different MFs along with three folds, whereby all yielded high to very high values, but the DsigMF showed the best result in CV. Therefore, we presented two evaluation methods: first, we used RMSE and R^2 with different folds of test data, and, second, we applied each fold of validation data to evaluate the results within the ROC curves, and DsigMF showed the best performance for the land subsidence susceptibility assessment in overall. To sum up, as ANFIS is a very technical method and sensitive to design decisions in all of the required steps, including the selection of the MFs, it is essential to choose the right parameters. At the same time, the model is less sensitive to differences in expert opinions. That is why this model is more reliable than the models that are based on expert knowledge. The limitation of this model is that researchers cannot directly obtain the absolute importance/weight of each individual factor.

Acknowledgements Open access funding provided by Austrian Science Fund (FWF). This research was partly funded by the Austrian Science Fund (FWF) through the Doctoral College GIScience (DK W 1237-N23) at the University of Salzburg. We would like to thank for support and help of Dr. Bakhtiar Feizizadeh (Department of Remote Sensing and GIS, University of Tabriz).

Author contributions Omid Ghorbanzadeh conceived the study, drafted most of the manuscript, performed the model and the statistical analysis. Thomas Blaschke structured the manuscript, helped with language formulations and wrote parts of the discussion and conclusion section. Hashem Rostamzadeh and Khalil Gholaminia supported the main author with the statistical and mathematical calculations. Jagannath Aryal

gave advice on the ANFIS modeling part and contributed in refining the validation section. All authors read and approved the final manuscript.

Compliance with ethical standards

Conflicts of interest The authors declare that they have no conflict of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Anwer N, Abbas A, Mazhar A, Hassan S (2012) Measuring weather prediction accuracy using sugeno based Adaptive Neuro Fuzzy Inference system, grid partitioning and guassmf. In: 2012 8th international conference on computing technology and information management (ICCTM), IEEE, vol 1, pp. 214–219
- Arlot S, Celisse A (2010) A survey of cross-validation procedures for model selection. *Stat Surv* 4:40–79
- Armaghani DJ, Hajihassani M, Sohaei H, Mohamad ET, Marto A, Motaghedi H, Moghaddam MR (2015a) Neuro-fuzzy technique to predict air-overpressure induced by blasting. *Arab J Geosci* 8(12):10937–10950
- Armaghani DJ, Hajihassani M, Monjezi M, Mohamad ET, Marto A, Moghaddam MR (2015b) Application of two intelligent systems in predicting environmental impacts of quarry blasting. *Arab J Geosci* 8(11):9647–9665
- Ashraf H, Cawood F (2015) Geospatial subsidence hazard modelling at Sterkfontein Caves. *South Afr J Geomat* 4(3):273–284
- Bardestani S, Givhechi M, Younesi E, Sajjadi S, Shamshirband S, Petkovic D (2017) Predicting turbulent flow friction coefficient using ANFIS technique. *SIViP* 11(2):341–347
- Barzegar R, Moghaddam AA, Tziritis E, Fakhri MS, Soltani S (2017) Identification of hydrogeochemical processes and pollution sources of groundwater resources in the Marand plain, northwest of Iran. *Environ Earth Sci* 76(7):297
- Basser H, Shamshirband S, Petković D, Karami H, Akib S, Jahangirzadeh A (2014) Adaptive neuro-fuzzy prediction of the optimum parameters of protective spur dike. *Nat Hazards* 73(3):1439–1449
- Boria RA, Olson LE, Goodman SM, Anderson RP (2014) Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. *Ecol Model* 275:73–77
- Bui DT, Pradhan B, Lofman O, Revhaug I, Dick OB (2012) Landslide susceptibility mapping at Hoa Binh province (Vietnam) using an adaptive neuro-fuzzy inference system and GIS. *Comput Geosci* 45:199–211
- Cabrera-Barona P, Ghorbanzadeh O (2018) Comparing classic and interval analytical hierarchy process methodologies for measuring area-level deprivation to analyze health inequalities. *Int J Environ Res Public Health* 15(1):140
- Cakit E, Karwowski W (2017) Predicting the occurrence of adverse events using an adaptive neuro-fuzzy inference system (ANFIS) approach with the help of ANFIS input selection. *Artif Intell Rev* 48(2):139–155
- Cam E, Yildiz O (2006) Prediction of wind speed and power in the central Anatolian region of Turkey by adaptive neuro-fuzzy inference systems (ANFIS). *Turk J Eng Environ Sci* 30(1):35–41
- Camastra F, Ciaramella A, Giovannelli V, Lener M, Rastelli V, Staiano A, Staiano G, Starace A (2015) A fuzzy decision system for genetically modified plant environmental risk assessment using Mamdani inference. *Expert Syst Appl* 42(3):1710–1716
- Cavallaro F (2015) A Takagi–Sugeno fuzzy inference system for developing a sustainability index of biomass. *Sustainability* 7(9):12359–12371
- Chai T, Draxler RR (2014) Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature. *Geosci Model Dev* 7(3):1247–1250
- Chen W, Pourghasemi HR, Zhao Z (2017a) A GIS-based comparative study of Dempster–Shafer, logistic regression and artificial neural network models for landslide susceptibility mapping. *Geocarto Intern* 32(4):367–385

- Chen W, Pourghasemi HR, Panahi M, Kornejady A, Wang J, Xie X, Cao S (2017b) Spatial prediction of landslide susceptibility using an adaptive neuro-fuzzy inference system combined with frequency ratio, generalized additive model, and support vector machine techniques. *Geomorphology* 297:69–85
- Dehghani M, Rastegarfar M, Ashrafi RA, Ghazipour N, Khorramrooz HR (2014) Interferometric SAR and geospatial techniques used for subsidence study in the Rafsanjan plain. *Am J Environ Eng* 4(2):32–40
- Dehghan-Soraki Y, Sharifikia M, Sahebi MR (2015) A comprehensive interferometric process for monitoring land deformation using ASAR and PALSAR satellite interferometric data. *GISci Remote Sens* 52(1):58–77
- Dixon B (2005) Applicability of neuro-fuzzy techniques in predicting ground-water vulnerability: a GIS-based sensitivity analysis. *J Hydrol* 309(1):17–38
- Erlacher C, Jankowski P, Blaschke T, Paulus G, Anders KH (2017) A GPU-based parallelization approach to conduct spatially-explicit uncertainty and sensitivity analysis in the application domain of landscape assessment. *GI_Forum* 2017(1):44–58
- Esen H, Esen M, Ozsolak O (2017) Modelling and experimental performance analysis of solar-assisted ground source heat pump system. *J Exp Theor Artif Intell* 29(1):1–17
- Feizizadeh B, Ghorbanzadeh O (2017) GIS-based interval pairwise comparison matrices as a novel approach for optimizing an analytical hierarchy process and multiple criteria weighting. *GI_Forum* 1:27–35
- Feizizadeh B, Kienberger S (2017) Spatially explicit sensitivity and uncertainty analysis for multicriteria-based vulnerability assessment. *J Environ Plan Manag* 60(11):2013–2035
- Feizizadeh B, Blaschke T, Roodposhti MS (2013) Integrating GIS-based fuzzy set theory in multicriteria evaluation methods for landslide susceptibility mapping. *Intern J Geoinform* 9(3):49–57
- Feizizadeh B, Roodposhti MS, Jankowski P, Blaschke T (2014a) A GIS-based extended fuzzy multi-criteria evaluation for landslide susceptibility mapping. *Comput Geosci* 73:208–221
- Feizizadeh B, Jankowski P, Blaschke T (2014b) A GIS based spatially-explicit sensitivity and uncertainty analysis approach for multi-criteria decision analysis. *Comput Geosci* 64:81–95
- Feizizadeh B, Roodposhti MS, Blaschke T, Aryal J (2017) Comparing GIS-based support vector machine kernel functions for landslide susceptibility mapping. *Arab J Geosci* 10(5):122
- Folorunsho JO, Igusi EO, Mu'azu MB, Garba S (2012) Application of adaptive neuro fuzzy inference system (Anfis) in River kaduna discharge forecasting. *Res J Appl Sci Eng Technol* 4(21):4275–4283
- Galloway DL, Jones DR, Ingebritsen SE (1999) Land subsidence in the United States, vol 1182. US Geological Survey
- Ganguli M (2011) Groundwater withdrawal and land subsidence: a study of Singur Block, West Bengal, India. *Intern J Geom Geosci* 2(2):465
- Gaspar JL, Goulart C, Queiroz G, Silveira D, Gomes A (2004) Dynamic structure and data sets of a GIS database for geological risk analysis in the Azores volcanic islands. *Nat Hazards Earth Syst Sci* 4(2):233–242
- Ghorbanzadeh O, Feizizadeh B, Blaschke T (2017) Multi-criteria risk evaluation by integrating an analytical network process approach into GIS-based sensitivity and uncertainty analyses. *Geomat Nat Hazards Risk* 9(1):127–151
- Gilks WR, Richardson S, Spiegelhalter D (eds) (1995) Markov chain Monte Carlo in practice. CRC Press, Boca Raton
- Hahn PR, Mukherjee S, Carvalho C (2010) Predictor-dependent shrinkage for linear regression via partial factor modeling. *arXiv preprint [arXiv:1011.3725](https://arxiv.org/abs/1011.3725)*
- Hajalilou B, Khaleghi F (2009) Investigation of hydro geochemical factors and groundwater quality assessment in Marand Municipality, northwest of Iran: a multivariate statistical approach. *J Food Agric Environ* 7(3&4):930–937
- Jang JS (1993) ANFIS: adaptive-network-based fuzzy inference system. *IEEE Trans Syst Man Cybern* 23(3):665–685
- Karimzadeh S (2015) Characterization of land subsidence in Tabriz (NW Iran) using watershed and InSAR analyses. *Acta Geodaetica Geophys*, Springer, 51: 181–195
- Karimzadeh S, Cakir Z, Osmanoglu B, Schmalzle G, Miyajima M, Amiraslazadeh R, Djamour Y (2013) Interseismic strain accumulation across the North Tabriz Fault (NW Iran) deduced from InSAR time series. *J Geodyn*, Elsevier, 66: 53–58
- Khaleghi F, Shahverdizadeh GH (2014) Hydrogeochemical characteristics and evaluation of the drinking and irrigation water quality in Marand Plain, East Azerbaijan, NW Iran. *Br J Appl Sci Technol* 4(17):2458
- Khorrami B (2016) Assessment of groundwater withdrawal applying GIS. M.Sc. Thesis, University of Tabriz, Tabriz, p. 110
- Kohavi R (1995) A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Ijcai*, vol. 14, no. 2, pp. 1137–1145)

- Lee S, Park I (2013) Application of decision tree model for the ground subsidence hazard mapping near abandoned underground coal mines. *J Environ Manage* 127:166–176
- Lee S, Ryu JH, Won JS, Park HJ (2004) Determination and application of the weights for landslide susceptibility mapping using an artificial neural network. *Eng Geol* 71(3):289–302
- Lee S, Park I, Choi JK (2012) Spatial prediction of ground subsidence susceptibility using an artificial neural network. *Environ Manage* 49(2):347–358
- Ligmann-Zielinska A, Jankowski P (2012) Impact of proximity-adjusted preferences on rank-order stability in geographical multicriteria decision analysis. *J Geogr Syst* 14(2):167–187
- Lim C, Yu B (2016) Estimation stability with cross-validation (ESCV). *J Comput Gr Stat* 25(2):464–492
- Lin CC, Lin CL, Shyu JZ, Lin CT (2012) The ANFIS system for nonlinear combined fore-casts in the telecommunications industry. *Intern J Comput Appl* 37(12):30–35
- Lombardo L, Cama M, Conoscenti C, Märker M, Rotigliano E (2015) Binary logistic regression versus stochastic gradient boosted decision trees in assessing landslide susceptibility for multiple-occurring landslide events: application to the 2009 storm event in Messina (Sicily, southern Italy). *Nat Hazards* 79(3):1621–1648
- MWREP: Ministry of water resource for East Azerbaijan Province (2014) Ground water depth assessment in Marand County. National project
- Naderloo L, Javadikia H, Mostafaei M (2017) Modeling the energy ratio and productivity of biodiesel with different reactor dimensions and ultrasonic power using ANFIS. *Renew Sustain Energy Rev* 70:56–64
- Navas JM, Telfer TC, Ross LG (2012) Separability indexes and accuracy of neuro-fuzzy classification in geographic information systems for assessment of coastal environmental vulnerability. *Ecol Inform* 12:43–49
- Nefeslioglu HA, Sezer EA, Gokceoglu C, Ayas Z (2013) A modified analytical hierarchy process (M-AHP) approach for decision support systems in natural hazard assessments. *Comput Geosci* 59:18
- Oh HJ, Pradhan B (2011) Application of a neuro-fuzzy model to landslide-susceptibility mapping for shallow landslides in a tropical hilly area. *Comput Geosci* 37(9):1264–1276
- Pacheco J, Arzate J, Rojas E, Arroyo M, Yutis V, Ochoa G (2006) Delimitation of ground failure zones due to land subsidence using gravity data and finite element modeling in the Querétaro valley, México. *Eng Geol* 84(3):143–160
- Pandey A, Sinha AK (2015) An empirical model of regional growth using adaptive neuro-fuzzy inference system. *Intern J Comput Appl*, 114(3)
- Parish ES, Kodra E, Steinhäuser K, Ganguly AR (2012) Estimating future global per capita water availability based on changes in climate and population. *Comput Geosci* 42:79–86
- Pirnazar M, Zand Karimi A, Feizizadeh B, Ostad-Ali-Askari K, Eslamian S, Hasheminasab H, Ghorbanzadeh O, Haeri Hamedani M (2017) Assessing flood hazard using GIS based multi-criteria decision making approach; study area: East Azerbaijan province (Kaleibar Chay basin). *J Flood Eng* 8(2):203–223
- Polykretis C, Chalkias C, Ferentinou M (2017) Adaptive neuro-fuzzy inference system (ANFIS) modeling for landslide susceptibility assessment in a Mediterranean hilly area. *Bull Eng Geol Environ*. <https://doi.org/10.1007/s10064-017-1125-1>
- Pourghasemi HR, Pradhan B, Gokceoglu C, Moezzi KD (2012) Landslide susceptibility mapping using a spatial multi criteria evaluation model at Haraz Watershed, Iran. In: *Terrigenous mass movements*, Springer, Berlin, pp. 23–49
- Pradhan B (2011) Use of GIS-based fuzzy logic relations and its cross application to produce landslide susceptibility maps in three test areas in Malaysia. *Environ Earth Sci* 63(2):329–349
- Pradhan B, Lee S (2010) Delineation of landslide hazard areas on Penang Island, Malaysia, by using frequency ratio, logistic regression, and artificial neural network models. *Environ Earth Sci* 60(5):1037–1054
- Şalap-Ayça S, Jankowski P (2016) Integrating local multi-criteria evaluation with spatially explicit uncertainty-sensitivity analysis. *Spat Cognit Comput* 16(2):106–132
- Sengur A (2008) Wavelet transform and adaptive neuro-fuzzy inference system for color texture classification. *Expert Syst Appl* 34(3):2120–2128
- Sezer EA, Pradhan B, Gokceoglu C (2011) Manifestation of an adaptive neuro-fuzzy model on landslide susceptibility mapping: Klang valley, Malaysia. *Expert Syst Appl* 38(7):8208–8219
- Shabankareh M, Hezarkhani A (2016) Copper potential mapping in Kerman copper bearing belt by using ANFIS method and the input evidential layer analysis. *Arab J Geosci* 9(5):364
- Shadman Roodposhti M, Aryal J, Shahabi H, Safarrad T (2016) Fuzzy Shannon entropy: a hybrid GIS-based landslide susceptibility mapping method. *Entropy* 18(10):343
- Singh R, Kainthola A, Singh TN (2012) Estimation of elastic constant of rocks using an ANFIS approach. *Appl Soft Comput* 12(1):40–45

- Singh R, Umrao RK, Ahmad M, Ansari MK, Sharma LK, Singh TN (2017) Prediction of geomechanical parameters using soft computing and multiple regression approach. *Measurement* 99:108–119
- Sun H, Grandstaff D, Shagam R (1999) Land subsidence due to groundwater withdrawal: potential damage of subsidence and sea level rise in southern New Jersey, USA. *Environ Geol* 37(4):290–296
- Swets JA (1988) Measuring the accuracy of diagnostic systems. *Science* 240(4857):1285–1293
- Tahmasebi P, Hezarkhani A (2012) A hybrid neural networks-fuzzy logic-genetic algorithm for grade estimation. *Comput Geosci* 42:18–27
- Takagi T, Sugeno M (1985) Fuzzy identification of systems and its applications to modeling and control. *IEEE Trans Syst Man Cybern* 1:116–132
- Umar Z, Pradhan B, Ahmad A, Jebur MN, Tehrany MS (2014) Earthquake induced landslide susceptibility mapping using an integrated ensemble frequency ratio and logistic regression models in West Sumatera Province, Indonesia. *CATENA* 118:124–135
- Václavík T, Meentemeyer RK (2009) Invasive species distribution modeling (iSDM): are absence data and dispersal constraints needed to predict actual distributions? *Ecol Model* 220(23):3248–3258
- VaeziNejad SM, Tofigh MM, Marandi SM (2011) Zonation and prediction of land subsidence (case study-Kerman, Iran). *Intern J Geosci* 2(02):102
- Vahidnia MH, Alesheikh AA, Alimohammadi A, Hosseinali F (2010) A GIS-based neuro-fuzzy procedure for integrating knowledge and data in landslide susceptibility mapping. *Comput Geosci* 36(9):1101–1114
- Wang FK, Chang KK, Tzeng CW (2011) Using adaptive network-based fuzzy inference system to forecast automobile sales. *Expert Syst Appl* 38(8):10587–10593
- Wiens TS, Dale BC, Boyce MS, Kershaw GP (2008) Three way k-fold cross-validation of resource selection functions. *Ecol Model* 212(3–4):244–255
- Fakhri M, Asghari Moghadam A, Najib M, Barzegar R (2015) Nitrate concentration in groundwater resources in Marand plain and the groundwater vulnerability assessment by AVI and GODS methods. In: *Proceedings of national conference of Ecology*, vol: 41, no. 1, Iran, pp. 49–66
- Zhang N, Xiao C, Liu B, Liang X (2017) Groundwater depth predictions by GSM, RBF, and ANFIS models: a comparative assessment. *Arab J Geosci* 10(8):189

Affiliations

Omid Ghorbanzadeh¹  · Hashem Rostamzadeh² · Thomas Blaschke¹ · Khalil Gholaminia³ · Jagannath Aryal^{4,5}

¹ Department of Geoinformatics–Z_GIS, University of Salzburg, Salzburg, Austria

² Department of Geography and Planning, University of Tabriz, Tabriz, Iran

³ Department of Remote Sensing and GIS, University of Tabriz, Tabriz, Iran

⁴ Visiting Academic, University of Salzburg, Salzburg, Austria

⁵ Discipline of Geography and Spatial Sciences, University of Tasmania, Hobart, Australia